DOCUMENT RESUME

ED 183 624 TM 800 097

AUTHOR Bormuth, John R.
TITLE Proficiency-Referencing a Reading Achievement Test.
PUB DATE Jan 79
NOTE 30p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Criterion Referenced Tests: Elementary Secondary Education: *Equivalency Tests: *Expectancy Tables: Mastery Tests: Models: Multiple Regression Analysis: *Norm Referenced Tests: Prediction: *Reading Tests: Scores: Scoring Formulas: Statistical Analysis: Tables (Data): Task Analysis: Test Construction: Testing Problems: *Test Interpretation: Test Reliability
IDENTIFIERS California Achievement Tests: Test Analysis

ABSTRACT

A procedure is demonstrated for constructing tables showing, for each score on a commercial reading achievement test, the percentage of real-world materials that the testee is likely to comprehend with at least a criterion level of proficiency, the percentages of students in a local or national sample who can competently comprehend a given percentage of materials, and which individuals or percentages of local and national samples have reached the state of being literate on the type of materials used to reference the test. Students are given a commercial reading test and some comprehensibility tests made from a sample of real-world reading tasks. The commercial test scores are regressed on the individual and mean comprehensibility test scores, and the regression equation is solved, fixing the comprehensibility score at the desired level of proficiency and varying the comprehensibility test mean to correspond to percentile points in the distribution of those means. The scores thus predicted are then used, along with the commercial test's norms, to construct the described tables. (Author/MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED183624

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PROFICIENCY-REFERENCING A READING ACHIEVEMENT TEST[1]

John R. Bormuth

University of Chicago

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
John R. Bormuth

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 800097

Running Head: Proficiency-Referencing a Reading Achievement Test

January, 1979

ABSTRACT

This report demonstrates a procedure for constructing tables showing, for each score on a commercial reading achievement test, the percentage of real-world materials that the testee is likely to comprehend with at least a criterion level of proficiency, the percentages of students in a local or national sample that can competently comprehend a given percentage of materials, and (at least conceivably) which individuals or percentages of local and national samples have reached the state of being "literate" on the type of materials used to reference the test. Students are given a commercial reading test and some comprehensibility tests made from a sample of real-world reading tasks. The commercial test scores are then regressed on the individual and mean comprehensibility test scores, and the regression equation is solved, fixing the comprehensibility score at the desired level of proficiency and varying the comprehensibility test mean to correspond to percentile points in the distribution of those means. The scores thus predicted are then used, along with the commercial test's norms to construct the tables described.

Proficiency-Referencing a Reading Achievement Test

An analysis of the tasks involved in managing educational resources generally indicates that decision makers should be supplied with at least three very different kinds of information that can, at least conceivably, be obtained using psychological tests. Elected and appointed officials need proficiency information, information telling them whether the knowledge students are learning is adequate for coping with the real-world tasks they are likely to encounter. These officials also need productivity information, information telling them whether the knowledge that some individual or group of students is getting compares favorably with that of other individuals or groups of students having the advantages of comparable personal, community, and school resources. And, third, teachers and curriculum workers need mastery information, information telling them whether and how well students are mastering individual skills -- the fragments and structures of

knowledge that are taught and assembled to form the end-product of education.

Commercial achievement tests often do a reasonably adequate job of furnishing information for mastery and productivity assessment. That is, scores on subtests and responses to individual items yield somewhat sparse but definitely useful estimates of the extent to which students master the skills being taught. And the norms on the tests, combined with other information, provide a useful basis for gauging whether the instruction is making comparatively effective use of a student's and a school's resources.

But these tests have stood mute on the crucial issue of proficiency. For all that they reveal, instruction may be imparting exactly enough knowledge to enable graduates to cope with the real-world tasks they are likely to encounter, or it may be wastefully teaching either far more or far less knowledge than they realistically want or need. To remedy this deficit, a number of efforts (e.g., Scott & Clark, 1976; Adult Performance Level Project, 1973, 1975; Murphy, 1973, 1975; Gadway & Wilson, 1973; Louis Harris and Associates, 1970, 1971) have been made to develop proficiency tests that ostensibly address the issue of what percentage of tasks a student can perform at a desirable level of proficiency. But the problems encountered were legion. Consider, as examples, just three of them.

One is that proficiency at performing a real-world task was commonly tested using questions that were not generated in a formally disciplined way, raising the issue of whether test writers biased their items by what content they chose to test and the way they phrased them. Then, since these informal test-making procedures allow a given score on these tests to correspond, willy-nilly, to almost any amount of the actual proficiency supposedly being tested, no credible criterion score could be set and used to count and report the fraction of tasks that a student could perform with desirable proficiency.

Another was that the two efforts made to avoid these problems rendered the tests irrelevant to the very issues they were supposed to address. One practice, reporting a count of the number or percent of items answered for all tasks combined, often fails to reveal whether a student was able to reach a desirable level of proficiency on even one task of any importance. And the other practice, reporting the percentage of students answering each individual item, obviously obscures that issue even more.

Finally, the plausibility of such tests depends on their containing a broad enough sample from a domain of tasks that the student's score can be regarded as a reliable estimate of the percentage of tasks, in the domain itself, that he could perform with the proficiency

desired. But testing him on that many tasks demands more testing time than schools feel is normally justified. And testing a student on only a couple of tasks and reporting the percentage of students reaching proficiency on each separate task, again, begs the issue supposedly being addressed.

The purpose of this report is to show that it is feasible to proficiency-reference a commercial reading achievement test: First, it describes how to develop a table showing for each score on that test a proficiency rate that represents the percentage of real-world materials that a student is likely to comprehend with at least a criterion level of proficiency. Second, it demonstrates the construction of a table showing proficiency rates of students attaining successively higher percentile scores among the group used to norm the test. And, third, it illustrates a method for assessing, with respect to a certain type of material, the literacy rate in the norm group, that is the percentage of students that is able to reach at least a criterion rate of proficiency regarded as separating the literate from the subliterate or the subliterate from the illiterate.

Rationale and Procedure

Major steps. This procedure can be seen as consisting of eight steps that begin with a) drawing a sample from a population of real-world tasks that students are

likely to want or need to read, b) using an algorithm to derive comprehensibility tests to measure proficiency in reading each of the materials, c) selecting a criterion score that represents the level of proficiency desired in comprehending materials of this sort, d) administering one of these proficiency tests and a commercial reading test to a sample of students, e) regressing the commercial test scores onto the proficiency test scores, and then f) using the regression equation and the criterion score to estimate the achievement test score that corresponds to the criterion score on this proficiency test.

This estimated score characterizes the difficulty of this particular material, expressing it in terms of the level of performance that a person would normally need to achieve on the commercial reading test in order to comprehend this material with at least the criterion level of proficiency. Then, by g) repeating these operations using the tests made from each of the other materials in the sample, we can build up a frequency distribution of predicted scores that characterizes the comprehensibilites of the materials in the parent population. And, since the predicted score of each material is now expressed in terms of scores on the commerical reading test, h) calculating percentiles of this distribution would produce a table that, in effect, proficiency-references the scores on this achievement test. That is, looking up a student's score

in this table would locate the percentile score which would represent the percentage of materials that he is likely to be able to comprehend with at least the criterion level of proficiency.

Data; The data used in this demonstration came from an earlier study (Bormuth, 1969) in which roughly 2,500 students were given the California Reading Test (1962) and (using a fractional factorial design) cloze tests made from 330 passages, each of which was drawn from a different material and all of which were commonly used in instruction. There were 33 passages representing each of the 10 subject matter categories, with 7, 8, 8, 7, and 3 of those passages being drawn from materials used primarily for instruction at the primary, elementary, junior high school, high school, and college levels respectively. Each passage was about 110 words in length, yielding cloze test forms of about 22 items each. The five forms of each cloze test were made by starting the deletion process with a different one of the first five words in a passage. The cloze tests were administered without time limits and without allowing students to read the intact passages and were scored by counting responses that, ignoring spelling errors, exactly matched th word deleted to form an item.

The students were enrolled in grades 4 through 12 in suburban, middle class schools, and were divided into 50 groups of about 57 students each using their California

test scores as the basis for a randomized blocking procedure. The students in each group took one of the forms of the cloze tests made from 33 passages which had been randomly assigned to the booklet assembled for that group. During a two week period that straddled the end of the fifth month of the school year, each elementary student took two booklets, while older students took only one. The scores for all five forms of a cloze test made from a passage were treated as if all were made on a single test, generally yielding 285, scores per passage.

Regression model. Instead of regressing the California Reading Test 330 different times, once for each cloze test, all of those regressions could be summarized in a single equation by exploiting the lawful relationships represented by the readability estimation model (Bormuth, 1978). This model regards ability to comprehend materials, c, as a complex trait composed partially of the reading skills, s, described in curricula and measured by commercial reading tests and partially of the verbal intelligence, i, needed to apply those skills to actual reading tasks. Roughly speaking, $c = s + i$. This notion derives from the long known fact (Bond & Tinker, 1967, p. 161; Harris, 1961, p. 180) that, to some extent, verbal intelligence and reading skill are mutually substitutable -- that, for example, a bright reader with few skills can comprehend a material as well as a less intelligent reader with more skills. For present purposes, let this principle be phrased as

$$s = c - i \quad . \qquad (1)$$

To operationalize this notion, s can be represented using a reader's, r, grade equivalent score on the California Reading Test, designating it with its capital letter, S. The term c can be only partially represented by the reader's score on a cloze test, C, made from a material. Since materials vary in ease, a given score could correspond to any number of levels of the underlying trait, c, unless C were adjusted using some measure of the material's ease, such as $\overline{C}$, the mean of the scores made on that cloze test. Thus, roughly speaking, c can be operationalized as $c = C - \overline{C}$. Finally, as will be explained next, the effects of a reader's intelligence can be creadibly and conveniently represented by conditioning the estimates of S on his grade level, G. That is to say,

$$S_r = f[C_{rm}, \overline{C}_r, (G_r)]. \qquad (2)$$

To understand this usage of G, consider first the analogy between S/G and the ratio in the definition $IQ = 100(MA/CA)$. Due to the common practice of granting automatic promotions in schools, $CA \simeq G + E$, where E is the average age at enrollment. Similarly, $MA \simeq S + E$ due to automatic promotions, due to the fact that MA and S are usually measured using very similar items, and due to the fact that S is expressed in the grade equivalent metric. Thus, MA/CA is substantively and mathematically akin to $(S + E)/(G + E)$ which in turn is a function of S/G.

Next, as indicated earlier, for a given level of $\underline{S}/\underline{G}$, $\underline{S}$ is estimated by

$$S(S/G) = \alpha_0 + \alpha_1\underline{C} - \alpha_2\overline{\underline{C}} \ . \qquad (3)$$

Therefore,

$$\underline{S}/\underline{G} = \alpha_0/\underline{G} + \alpha_1\underline{C}/\underline{G} - \alpha_2\overline{\underline{C}}/\underline{G} \ . \qquad (4)$$

Now, returning to Equation 1, substituting $\underline{S}$ for $\underline{s}$, and the right hand members of Equations 3 and 4, respectively, for $\underline{c}$ and $\underline{i}$ yields essentially the model fit to the data analyzed here, namely

$$\underline{S} = \alpha_0 + \alpha_1\underline{C} - \alpha_2\overline{\underline{C}} - (\alpha_0/\underline{G} + \alpha_1\underline{C}/\underline{G} - \alpha_2\overline{\underline{C}}/\underline{G}) \ . \qquad (5)$$

One variation from this model was that the sizes of the parameters were not restricted as shown here. Another, that added nothing, was to examine all possible interactions with the intention of retaining all that were significant. And a third was to obtain a better fit to the data by using the logit, $.5[\ln(\underline{p}1(1 - \underline{p})]$, of the cloze scores and by using raw grade level itself rather than its reciprocal, as Equation 5 specifies. A consequence of this last variation, as will be seen in Table 1, was to reverse the sign in front of the parenthesis in Equation 5.

The data were analyzed, first, using the step-wise regression procedure in order to identify the independent variables that would appear in the final equation and to obtain an initial estimate of their slopes and intercept. Terms were admitted and

retained at this stage only if they were significant at the .01 level. Next, following the procedures described by Mosteller and Tukey (1977), the data were broken down by passage into 33 subsets, each of which deleted a different, randomly chosen, set of 33 passages from the total set. Then, using each of these subsets successively, parameters were estimated for the terms retained in the original analysis; their deviations from the original estimates were calculated; and jackknifed estimates of the parameters and their standard errors were calculated using these estimates and their deviations. Table 1 shows these estimates and their *t*-ratios along with the correlations obtained in the first stage of this analysis. Although the jackknifed estimates of the standard errors were somewhat larger than those obtained at the first stage of the analysis, the jackknifed *t* ratios for all terms comfortably exceeded the value needed to reach the .01 level of significance, namely 2.59.

INSERT TABLE 1 ABOUT HERE

Although the regression equation obtained had a multiple correlation of only .885, it seems to have accounted for nearly all the reliable variance in this set of variables. This judgment was reached, in part, by noting that grade level was measured without error while the California and individual cloze tests generally had correlations very close to 1.00 when corrected for attenuation due to their reliabilities (Bormuth, 1969, p. 54). Also, a Rasch cloze score was calculated for each student (Wright, 1977), which used all the information contained in the cloze scores and the cloze means but did not entail the

assumption that the regression equation did, that the slopes of the regressions of the California test on the cloze tests were lawfully regular. Then, when the students' California test scores were regressed on their grade levels and Rasch cloze scores, the multiple correlation, .881, was only negligibly different.

Although this paper is not directly concerned with readability measurement theory, it is relevant here to demonstrate that the regression equation is creadible in terms of that theory. First, according to Equations 3 and 5, the main effects of cloze scores and cloze means account for the "effects" that ability to comprehend materials in general (c in Equation 1) have on reading skills. Their slopes were about equal and opposite, and their interaction did not enter the regression. This indicated, first, that variations in the trait underlying these two measures effect variations in both by about the same amount and in the same direction and, second, that, when intelligence is held constant, each level of reading skill is associated with exactly one level of the trait underlying the cloze scores and means. Second, according to Equations 4 and 5, the main effects and interactions involving grade level simply adjust the estimates of the level of a reader's skill to compensate for the role that intelligence plays in the composition of ability to comprehend materials in general. These terms did enter the regression equation, as expected. Furthermore, the slopes of those interactions were roughly equal but opposite, indicating that the size of these

adjustments was about constant for each level of the materials comprehension ability underlying the cloze scores and means.

Criterion score. Although domain-referenced, algorithmically-derived proficiency tests have the properties needed to identify rationally a criterion score that represents the most desirable level of performance (Bormuth, 1978) using the methods developed in utility theory (Fishburn, 1968) and decision analysis (Kenney & Raiffa, 1976), those identified so far (Bormuth, 1971) are fairly tenuous. However, the evidence available does provide the basis for identifying a criterion score that plausibly separates clearly incompetent performances from those that are at least marginally competent. That is, students who answer less than 35 percent of the items on a cloze test have roughly a 50-50 chance of gaining no information from reading a material, where information gained is measured as a residualized difference score obtained, in essence, by administering a 20 question comprehension test to a student before and after he has read a passage.

Calculating proficiency levels. The regression obtained was then solved separately for students within each of the grades, 4 through 12, by setting the cloze score, C, at the criterion level of .35 and then solving repeatedly for each of several selected values of cloze mean, $\overline{C}$. The values selected for the cloze means were obtained by plotting the cumulative frequency of passages whose means fell at or below each cloze score, fairing in a curve to smooth the minor

variations in this plot, and then using the smoothed curve to determine the cloze means corresponding to selected percentile points. These sets of solutions were plotted and then recast in the tabular form shown in Table 2.

Demonstrations

Table 2 shows the proficiency rates of students falling within the range of these data. The first column gives the grade-equivalent scores and the remaining columns show the proficiency rates at each grade level.

INSERT TABLE 2 ABOUT HERE

Perhaps the most striking feature of this table is the very different proficiency rates achieved by students who differed in grade level but not in grade equivalent score. As suggested above, this arose primarily because of the information that his score and grade level carry about the verbal intelligence of the student. A student in grade 4.5 with a score of 6.0, for example, would be able to call on not only that amount of reading skill but also on the superior IQ of $100(6.0 \div 4.5) = 133$, the combined effect of which would enable hi... to comprehend proficiently 81 percent of the materials. But a student in grade 5.5 having only that score would have a verbal IQ of only 109, leaving him able to apply that level of skill proficiently to only 14 percent of these materials.

Failing to condition the equation on grade level would have led to clinically preposterous estimates, to the estimate that, for example, a score of 9.2 was required of anyone to reach proficiency on a material with a cloze mean of .39, which was typical of the news feature articles included in this study. This implies that only a rare genious (IQ=202) in grade 4.5 could comprehend this material proficiently, while even a mental retardate (IQ=74) in grade 12.5 could do so. As the equation stands, it yields the more creadible estimates that students in grades 4.5 and 12.5 could reach proficiency on this material if they have grade equivalent scores of 6.0 and 12.5, respectively, implying IQ's of 133 and 102, respectively.

The data in Table 2 are useful primarily to those who wish to evaluate raw and grade-equivalent scores on a commercial reading test. But those wishing to evaluate the adequacy of a student's instruction would probably find it more useful to know what percentage of students was reaching each proficiency rate. This could be determined either by using students' percentile scores on the commercial reading test as the dependent variable in the regression model (recall that the predictions are conditioned on grade level) or by, as was the case here, simply looking up the percentile scores in the norm tables accompanying the test. The results for the twelfth graders are shown in Table 3. If the test is scored by a properly equipped scoring service, these proficiency rates could also be easily determined for percentile

INSERT TABLE 3 ABOUT HERE

points in the distributions of scores obtained for students within specific schools and school districts.

Essentially this same procedure could conceivably be used to assess the literacy of a student or the literacy rate in a group, where a student is regarded as being literate if he can comprehend with at least a criterion level of proficiency a criterion proportion of the materials that he is likely to want or need to read. This type of assessment cannot presently be demonstrated directly since we do not yet have a fully credible criterion of proficient comprehension of single materials much less a literacy criterion that, all things considered, marks the most desirable fraction of materials for a person to comprehend proficiently. At best, we might presently try to use our 35 percent criterion to distinguish those who are incompetent with respect to a material from those who are at least marginally proficient and then construct a table, such as Table 4, that allows a reader to select his own "literacy" criterion. Thus, if one wants to regard a twelfth grade student as being at least marginally literate when he can comprehend with at least marginal proficiency, say, at least 90 percent

INSERT TABLE 4 ABOUT HERE

of the type of materials tested, we could note from Table 3 that he would need a grade equivalent score of about 12.9. And then by rearranging data of that sort we obtain Table 4 which

shows that only 45 percent of the twelfth grade students in the norm group were at least marginally literate. And this, of course, implies that 100 - 45 = 55 percent were illiterate by these standards relative to materials of the type tested.

Discussion

While these results seem adequate to demonstrate the nature of this technique for proficiency-referencing commercial achievement tests, several cautions and amplifications may be useful to consider.

The materials used in this demonstration are less than ideally informative, since they are not very representative of those that a specific individual or group might realistically want or need to read at present or in the near or distant future for which he is preparing. Results from an actual application of this technique would be far more informative if readers were stratified by age, ability, aspirations, and the like and then their proficiency rates were reported for just those materials that are most appropriate for each group of readers, stratifying and weighting those materials also to correspond to the kinds and amounts of uses likely to be made of them.

As the procedure was described, it may be needlessly expensive and inflexible, since it requires the direct testing of a large number of students using a large number of materials. However, measures of the linguistic complexity of a material could be used in Equation 5 to replace the term $\hat{C}$, which represents the difficulty of a material. Doing so

would yield a generalized version of a readability formula. In another study (in progress) the author found that such formulas have nearly as much accuracy as the equation derived on the pattern of Equation 5. Therefore, one could effect major economies with almost no loss of accuracy by developing one such formula and then using it on various samples of materials to build up proficiency-referencing tables of the sorts either illustrated or suggested here. Ideally, the testing of the materials needed to develop such formulas would be conducted (using matrix sampling designs for assigning comprehensibility tests to students) as a routine part of the norming studies conducted when a publisher develops or updates a test.

Finally, it is problematic if it is presently feasible to use this technique to proficiency-reference commercial achievement tests in other content areas. For one thing, doing so requires that we be able to identify and delimit real-world tasks to which the content applies. While it is fairly clear that a newspaper editorial, a book and the directions on a package represent reading tasks, it is less than obvious what might be the limits of a task for applying knowledge of mathematics, history, chemistry, or the like. For another, doing so requires the use of an algorithmically generated proficiency test. And those who study instruction in such content areas have given the problem little attention. And, for yet another thing, the performance of such tasks is more likely to call upon the application of knowledge from several

content areas rather than merely knowledge from one, requiring the use of a canonical regression model rather than the simple one described in Equation 5.

The procedure needed to proficiency-reference a commercial test of writing composition ability may be somewhat more analogous to the one described here, since a writer's literacy, in an important sense of that slippery term, refers to his ability to write prose that is comprehensible to the readers to whom he wants or needs to communicate information. But his ability to do so depends not only on his ability to adapt his writings to suit the motives and skills of his readers but also on social and economic factors that impose brevity and a number of other constraints. So even this analogy seems unlikely to be as simple and direct as it may at first appear.

In any case, this technique does appear to be useful for proficiency-referencing commercial reading tests, and that single application is by itself sufficiently important to warrant the efforts needed to further develop and apply the technique.

REFERENCES

Adult Performance Level Project. Adult functional competency. Austin, Texas: University of Texas Extension, 1975a.

Adult Performance Level Project. Adult functional competency: A summary. Austin, Texas: University of Texas Extension, 1975b.

Adult Performance Level Project. The adult performance level study. Austin, Texas: University of Texas Extension, 1973.

Bond, G.L., & Tinker, M.A. Reading difficulties: Their diagnosis and correction. New York: Appleton-Century-Crofts, 1967.

Bormuth, J.R. Feasibility of criterion-referenced testing: Regression identity hypothesis. Paper presented at the meeting of the American Educational Research Association, Toronto, 1978.

Bormuth, J.R. Development of standards of readability: Toward a rational criterion of passage performance (U.S.O.E. Tech. Rep. 9-0237). Chicago: University of Chicago, 1971.

Bormuth, J.R. Development of readability analyses (U.S.O.E. Tech. Rep. 7-0052). Chicago: University of Chicago, 1969.

Bormuth, J.R. Cloze readability criterion reference scores. Journal of Educational Measurement, 1968, 5, 189-196.

Bormuth, J.R. Comparable cloze and multiple-choice comprehension scores. Journal of Reading, 1967, 10, 291-299.

Clark, J.P. & Scott, D.T. Competency tests and graduation requirements. Reston, Virginia: National Association of Secondary Principals, 1976.

Fishburn, P.C. Utility theory. Management Science, 1968, 14, 335-378.

Gadway, C., & Wilson, H.A. Functional literacy: Basic reading performance. Denver: National Assessment of Educational Progress, 1973.

Harris, Louis & Associates. Survival literacy study (Study No. 2036). New York: Author, 1970.

Harris, Louis & Associates. The 1971 national reading difficulty index. New York: Author, 1971.

Keeney, R.L., & Raiffa, H. Decision making with multiple objectives: Preference and value tradeoffs. New York: Wiley, 1976.

Mosteller, F., & Tukey, J.W. Data analysis and regression. Reading, Mass.: Addison-Wesley, 1977.

Murphy, R.T. Adult functional reading study (73-48). Princeton: Educational Testing Service, 1973.

Murphy, R.T. Adult functional literacy study (PR 75-2). Princeton: Educational Testing Service, 1973.

Wright, B.D. Solving measurement problems with the Rasch model. Journal of Educational Measurement, 1977, 14, 97-116.

FOOTNOTES

1. Credit is due to Edward Haertle for programming the jackknifing analyses used in this study and to Stanley Sapahia for adapting it for use with this set of data and for performing the Rasch analysis used to indicate the sufficiency of the model employed here.

TABLE 1

Regression Analysis

Variables	Regression[a]		Correlations[b]						
	Coefficients	t	Cloze score	Cloze mean	Grade level	G x C	G x $\bar{C}$	Mean[b]	S. D.[b]
Reading skill =			.571	.001	.852	.422	-.241	9.51	3.06
Cloze score	1.13	28.99		.504	.465	.931	.329	-.292	.566
Cloze means	-1.25	-16.27			.000	.554	.918	-.292	.285
Grade level	.841	219.00				.283	-.284	8.52	2.56
G x C	.0843	13.11					.471	-1.82	4.37
G x $\bar{C}$	-.0690	-6.79						-2.49	2.64
Intercept	2.29	75.34							

[a]Estimates obtained using jackknifing procedures.

[b]Based on all data.

Table 2

Percentages of Material that Students Can Comprehend With at Least the Criterion Level of Proficiency

Grade Equivalent Score	Student's grade level in school								
	4	5	6	7	8	9	10	11	12
5.0	3								
5.5	36	<1							
6.0	81	14	<1						
6.5	98	53	2						
7.0	>99	90	28	<1					
7.5		>99	69	9	<1				
8.0			95	42	1				
8.5			>99	81	21	<1			
9.0				98	57	6	<1		
9.5				>99	89	34	2		
10.0					>99	71	15	<1	
10.5						94	46	3	
11.0						>99	82	37	<1
11.5							97	62	12
12.0							>99	89	39
12.5								98	73
13.0								>99	94
13.5									>99

Table 3

Grade Group Percentiles and Proficiency Rates Corresponding to Grade Equivalent Scores of Twelfth Grade Students

Grade equivalent scores	Percentiles of	
	students	materials
11.0	22	<1
11.1	24	1
11.2	25	2
11.3	26	3
11.4	28	8
11.5	30	12
11.6	31	16
11.7	34	22
11.8	35	28
11.9	37	33
12.0	38	39
12.1	40	45
12.2	42	53
12.3	44	60
12.4	46	68
12.5	48	73
12.6	50	79
12.7	52	84
12.8	53	88
12.9	55	92
13.0	57	94
13.1	59	96
13.2	61	97
13.3	63	98
13.4	65	99
13.5	67	>99

Table 4

Percentages of Students Classified as being at Least Marginally Literate or Illiterate on the Type of Materials Tested, Depending on the Proficiency Rate Chosen as the Criterion of Illiteracy

Criterion percent of materials	Grade								
	4	5	6	7	8	9	10	11	12
At least marginally literate:									
10	37	45	53	62	63	66	66	67	71
25	32	39	46	53	58	61	63	64	66
50	22	31	37	46	50	53	55	56	58
75	16	17	27	39	43	45	49	50	51
90	13	14	16	34	35	36	43	44	45
Illiterate:									
10	63	55	47	38	37	34	34	33	29
25	68	61	54	47	42	39	37	36	34
50	78	69	63	54	50	47	45	44	42
75	84	83	73	61	57	55	51	50	49
90	87	86	84	66	65	64	57	56	55